Kosslyn quarterly report
August 1991

HARVARD UNIVERSITY
DEPARTMENT OF PSYCHOLOGY

WILLIAM JAMES HALL
33 KIRKLAND STREET
CAMBRIDGE, MASSACHUSETTS 02138

8 August 1991

Dr. Terry Allard
Cognitive Science Programs
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Dear Terry:

I am writing this letter to inform you about progress on our project, "PET Studies of Components of High-Level Vision" (N00014-91-J-1243). We have made progress on two fronts during the last quarter.

I. PET Studies

We have finished one PET study and are about to begin two more, as noted below.

Primary visual cortex activation

We have finished analyzing the data from our first PET experiments and written a report of the results. As noted earlier, we have strong evidence (p = .0001) that primary visual cortex is activated selectively during visual mental imagery. We will mail you the paper (along with the slides) shortly.

Canonical and noncanonical views during object identification

In our next PET experiment we will study how objects are identified when seen from unusual points of view. Warrington and Taylor (1973) found that patients with right-parietal lesions have a very difficult time recognizing objects seen from unusual points of view. Kosslyn, Flynn, Amsterdam and Wang (1990) explain this result by positing a top-down, hypothesis testing mechanism that is called into play when a stimulus does not immediately match a stored memory very well. This mechanism not only relies on processes in the parietal lobe to shift attention, but also on processes in the frontal lobe to formulate hypotheses. To test these ideas, subjects will see a series of pictures, either of objects seen from a canonical point of view or of objects seen from an unusual point of view. One sec after seeing a picture, a word will be presented, and the subject will decide whether the word names the pictured object. Counterbalancing will ensure that the same objects and words appear equally often in the two conditions. In addition, in a baseline condition subjects will see random noise masks and hear names of objects, and will simply press a pedal when they hear the word. By subtracting the blood flow evoked by this baseline task from that evoked when canonical pictures are presented, we can examine the brain bases of bottom-up picture naming; by subtracting the blood flow evoked when canonical pictures are named from that

evoked when noncanonical pictures are named, we can examine the brain bases of top-down hypothesis testing.

Levels of object identification

Immediately after the next PET study, we will study the processes that underlie one's ability to name objects at different levels of hierarchy (these studies were summarized in the previous quarterly report). We have now selected a set of 48 objects that can be named at a subordinate, "basic," or superordinate level (e.g., one can name a rocking chair as *furniture*, *chair*, or *rocking chair*). This turned out to be a challenging task, and took more time than anticipated. We are now preparing an experiment to examine processing when subjects verify names at the different levels. We expect memory search to be used when one evaluates a superordinate name, if in fact one spontaneously labels pictures at the basic level (as has often been argued; see Kosslyn & Chabris, 1990, for a review). In this case, one must search memory for the superordinates of the object. In addition, we expect top-down processes to be used when one hears a subordinate name, which requires one to encode more visual information than what is needed to identify objects at the basic level. For example, a canary will spontaneously be identified as a "bird," and additional information (e.g., about its color, size, and specific shape) is necessary to affirm that it is a canary. Thus, the same areas that underlie naming noncanonical perspectives should also be involved in naming objects at a subordinate level. Moreover, these same top-down processes may be involved in visual mental imagery, as noted below.

II. Off-line Preliminary Studies

Because PET studies are so expensive and demanding, we typically perform off-line experiments as a prelude to a subsequent PET study. We have performed two such studies in an effort to ensure that the PET study is well-motivated and to develop appropriate tasks.

Two kinds of visual imagery

We have performed two off-line experiments to explore the idea that there are two types of visual imagery, parietal-based "attentional imagery" and temporal-based "visual memory imagery." The first sort of imagery relies on allocating attention selectively (as occurs when one visually "picks out" individual tiles on a tiled floor to form a pattern), and the second relies on activating a stored visual memory (as occurs when one visualizes one's mother's face). And in fact, in our previous PET study of imagery we found areas of the brain involved in attention (the pulvinar and the anterior cingulate) to be active when subjects formed images in grids.

In one experiment, subjects were tested in two conditions. In one, they formed images of letters in grids with their eyes open. This task requires selectively attending to specific rows and columns in the grid, and hence should reflect attentional imagery. In the other task, the subjects formed images of letters that were not in grids and had their eyes closed. This task should evoke visual-memory imagery. We found a larger difference in the time to form images of simple vs. complex images in the first condition than in the second. We expected such a larger effect of the number of segments when subjects use attentional imagery because each segment should be stored as a description of where to attend; the "attention window" is restricted to one location at a time, and hence must be moved sequentially when forming images that have multiple segments (previous work has supported this conjecture; e.g., Kosslyn et al., 1988). In contrast, when subjects use visual memory imagery, segments can be organized into higher-order perceptual units, and so fewer units need to be activated to generate the image--resulting in a diminished effect of complexity.

In another experiment, subjects were asked to listen to a sequence of directions (e.g., "north, northeast, west...") and to image a one-inch line segment pointing in each direction, with each

segment connected to the previous one. The segments were to form a pathway, tracing out the directions; the pathways included 5, 6, or 7 segments. At the end of the sequence was a beep; when the subjects heard it, they were to indicate whether the terminus of the pathway was above or below the starting point. We expected this task to rely on allocating attention over the visual buffer, and hence each segment would be represented separately (for the reasons noted above). In the second task, the subjects first were shown an array of dots on a piece of cardboard; the dots formed rows and columns that were unevenly spaced. Subjects again heard a list of directions, but now were told that the directions indicate how thick black lines (which were shown to them) connect up pairs of dots, starting at the one in the center. Again, when hearing the beep, they were to indicate whether the end point was above or below the beginning point. Because the rows and columns were unevenly spaced, the subjects were told that they must image the array to perform the task properly. Thus, we expected this task to require activating temporal-lobe based visual memories. We expected patterns of lines to be grouped into higher-order units, and hence images of these stimuli should be maintained more easily than attention-based images of the pathways in the first condition.

We compared the decision times, and found that the attention condition did in fact require more time. More interesting, however, was the finding that progressively more errors were committed in the attention condition with more complex pathways, which was not true in the visual-memory condition. This was as expected if the segments must be maintained separately in attention-based images, but can be grouped in visual-memory based images. In addition, we asked the subjects to rate various qualities of their images after each task, and found that the subjects rated the visual memory images as more vivid, sharper, and less like simply "paying attention" to a region. These data provide convergent evidence for the distinction between the two kinds of imagery.

We presently are conducting a follow-up study in which subjects must build up an image of a pathway, and then open their eyes to classify a pattern, and only then evaluate the image. We expect the intervening perceptual task to obliterate the image, which will require the subjects to generate the image anew when the probe is provided. Thus, we can measure the time subjects require to generate the two kinds of images. We expect a larger effect of the number of segments in the attentional imagery condition, for the same reasons noted above.

We plan to use both the letter-generation and path-formation tasks in future PET studies, which will provide convergent evidence for the distinction between the two kinds of imagery as well as information about the brain bases of these processes. We expect that visual memory imagery should evoke most of the same areas activated when top-down hypothesis testing is used during visual object identification, as noted above.

In short, this research is progressing on schedule, and we again thank you for your support.

Sincerely,

Stephen M. Kosslyn
Professor

References

Kosslyn, S. M., Cave, C. B., Provost, D., & Von Gierke, S. (1988). Sequential processes in image generation. *Cognitive Psychology, 20*(319-343),

Kosslyn, S. M., and Chabris, C. F. (1990). Naming pictures. *Journal of Visual Languages and Computing,, 1*, 77-96.

Kosslyn, S. M., Flynn, R. A., Amsterdam, J. B., & Wang, G. (1990). Components of high-level vision: A cognitive neuroscience analysis and accounts of neurological syndromes. *Cognition, 34*, 203-277.

Warrington, E. K., & Taylor, A. M. (1973). The contribution of the right parietal lobe to object recognition. *Cortex, 9*, 152-164.